Patent Application
MJM Do. No. 6647-17
Novell IDR-451

POLICY ENFORCEMENT USING THE SEMANTIC CHARACTERIZATION OF TRAFFIC

Sub A7

RELATED APPLICATION DATA

This application is related to U.S. Patent No. 6,108,619, issued August 22, 2000, titled "METHOD AND APPARATUS FOR SEMANTIC CHARACTERIZATION," to co-pending U.S. Patent Application Serial No. 09/512,963, titled "CONSTRUCTION, MANIPULATION, AND COMPARISON OF A MULTI-DIMENSIONAL SEMANTIC SPACE," filed February 25, 2000, to co-pending U.S. Patent Application Serial No. _________, titled "A METHOD AND MECHANISM FOR THE CREATION, MAINTENANCE, AND COMPARISON OF SEMANTIC ABSTRACTS," filed __________, and to U.S. Patent Application Serial No. __________, titled "INTENTIONAL-STANCE CHARACTERIZATION OF A GENERAL CONTENT STREAM OR REPOSITORY," filed simultaneously herewith and incorporated by reference herein, all commonly assigned.

FIELD OF THE INVENTION

This invention pertains to enforcing network policy decisions in a computer system, and more particularly to enforcing policy decisions by monitoring network traffic and content.

BACKGROUND OF THE INVENTION

In the prior art, policy enforcement is performed by counting packets traveling from their source to their destination. Most policy enforcement implementations ignore the content of the packets traversing the system. If the amount of traffic between source and destination Internet Protocol (IP) addresses becomes excessive, the policy enforcement implementation applies a limit to the packet flow.

One policy enforcement implementation (Layer 7) ostensibly considers the semantic content of the packets crossing the system. Layer 7 looks at tags in the header of the packet. If too many packets having a particular tag are crossing the system, Layer 7 restricts the flow of packets. But Layer 7 only considers tags in the packet header, and does not actually look at the semantic content of the packets. Thus, a program that sought to bypass the policy enforcement of Layer 7 only has to fraudulently label the tag in the header of the packet, and the policy will not be enforced against the packet.

Sub A2

U.S. Patent Application Serial No. __________, titled "INTENTIONAL-STANCE CHARACTERIZATION OF A GENERAL CONTENT STREAM OR REPOSITORY," filed simultaneously herewith, incorporated by reference herein, and referred to as "the Intentional Stance application," describes how users can listen to a content stream and set up response actions according to the content. Templates that include a set of state vectors in a topological vector space define the trigger. When the semantic content of the content stream comes close enough to the template, the action is triggered. But the Intentional Stance application does not describe how a network policy can be enforced using templates.

The present invention addresses these and other problems associated with the prior art.

SUMMARY OF THE INVENTION

The invention is a method and apparatus for enforcing policy over a computer network. A template is defined and assigned a policy. The network is then monitored to watch content in a content stream. When the content stream comes within a threshold distance of the template, the policy is enforced.

The foregoing and other features, objects, and advantages of the invention will become more readily apparent from the following detailed description, which proceeds with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows a computer system on which the invention can operate to use a template to enforce network policy.

FIG. 2 shows a two-dimensional topological vector space in which a template is presented.

FIG. 3 shows a two-dimensional topological vector space in which the distance between a template and a content stream is measured.

FIG. 4 shows a flowchart of a method according to the preferred embodiment of the invention to use a template in the computer system of FIG. 1 to enforce network policy.

FIG. 5 shows a flowchart of a method according to an alternate embodiment of the invention to use a template in the computer system of FIG. 1 to enforce network policy.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Sub B3

U.S. Patent Application Serial No. __________, titled "INTENTIONAL-STANCE CHARACTERIZATION OF A GENERAL CONTENT STREAM OR REPOSITORY," filed __________, incorporated by reference herein and referred to as "the Intentional Stance application," describes the creation of templates to use in intentional stance characterization. Recall that a template is a set of vectors defined by a semantic content. The template is compared with a content source. If the content source is close enough to the template, an action associated with the template is triggered.

FIG. 1 shows a computer system 105 on which templates can be used to enforce network policy. Computer system 105 conventionally includes a computer 110, a monitor 115, a keyboard 120, and a mouse 125. But computer system 105 can also be an Internet appliance, lacking monitor 115, keyboard 120, or mouse 125. Optional equipment not shown in FIG. 1 can include a printer and other input/output devices. Also not shown in FIG. 1 are the conventional internal components of computer system 105: e.g., a central processing unit, memory, file system, etc.

Computer system 105 further includes software 130. In FIG. 1, software 130 includes template 135, policy 140, network monitoring software 145, and policy enforcer 150. Template 135 is the template indicating the content to which associated policy 140 is to be applied. Network monitoring software 145 is responsible for monitoring the network, both for content and for metadata about the content. (Metadata is data about data, in this case about the content stream. For example, metadata about the content stream can include the percentage of traffic related to content close to the template. The collecting of metadata is represented pictographically by sensor 147.) Policy enforcer 150 is responsible for enforcing policy 140 when the content stream is close to template 135.

Although the content compared with template 135 can be found stored on computer system 105, this is not required. FIG. 1 shows computer system 105 accessing content stream 160 over network connection 165. Content stream 160 can be a single document, or it can include multiple sources. Content streams with multiple sources are common in today's world. For example, newsgroups and discussion lists (e-mail lists) allow multiple users to carry on several conversations on multiple topics at the same time. Newsgroups and discussion lists are typically organized into a hierarchy. The newsgroup itself has a content focus. This content focus is divided into sub-topics, called threads. Each thread is further divided into individual messages from individual users. FIG. 1 shows content stream 160 as having two threads. Thread one has two messages, and thread two has three messages.

Network connection 165 can include any kind of network connection. For example, network connection 165 can enable computer system 105 to access content stream 160 over a local area network (LAN), a wide area network (WAN), a global internetwork, or any other type of network. Similarly, once collected, the impact summary can be stored somewhere on computer system 105, or can be stored elsewhere using network connection 165.

FIG. 2 shows a two-dimensional topological vector space in which a template includes a set of state vectors. (FIGs. 2 and 3, although accurate representations of a topological vector space, are greatly simplified for example purposes, since most topological vector spaces will have significantly higher dimensions.) In FIG. 2, template 205 includes the state vectors represented by the "x" symbols. (For clarity, the line segments from the origin of the topological vector space to the heads of the state vectors are not shown in FIG. 2.)

Circle 210 represents the threshold distance defined for template 205 before the policy is enforced. The reader will recognize that circle 210 is an abstraction, since in the preferred embodiment distance is not measured from a single point in the topological vector space. Instead, in the preferred embodiment distance is measured from the entire set of vectors comprising the template, using the Hausdorff distance function or alternative measures suggested in the Intentional Stance application. But if template 205 could be reduced to a single point in the topological vector space, circle 210 could represent a threshold distance. Any content that comes within circle 210 would then trigger the policy associated with template 205.

FIG. 3 shows a two-dimensional topological vector space in which template 205 is compared with an impact summary for a content source. (To avoid clutter in the drawing, FIG. 3 shows template 205 and impact summary 305 in different graphs of the same topological vector space. The reader can imagine the template and impact summary as being in the same graph.) Using the Hausdorff or other distance function, the distance 310 between template 205 and impact summary 305 can be quantified. If distance 310 is smaller than the threshold distance defined for template 205, then the policy associated with template 205 will be triggered.

FIG. 4 shows a flowchart of a method according to the preferred embodiment of the invention to use a template to enforce network policy in the computer system of FIG. 1. At step 405 a template is defined. At step 410, a policy is defined and assigned to the template. At step 415, the content stream is monitored to see how close it comes to the template. At step 420, the network is monitored to determine metadata about the content stream (e.g., the percentage of network traffic devoted to the content stream triggering the template). At step 425, if the content stream comes close enough to be within the threshold distance for the template, the associated policy is enforced.

As an example of a possible network policy and its use, consider a server supporting newsgroup traffic. (As the reader will recall, a newsgroup carries multiple threads, each thread composed of messages generated by readers of the newsgroup.) One such newsgroup can be dedicated to medicine. Because the subject of abortion is generally controversial, the system administrator for the server can set a low bandwidth limit to messages relating to abortion. For example, the system administrator can set the policy to limit total bandwidth to messages relating to abortion at 5% of the bandwidth for the newsgroup. Similarly, the system administrator can set a policy dedicating a minimum guaranteed bandwidth to a subject, so that topical subjects are not lost for lack of bandwidth.

As an additional example, consider a network where security is an issue. In such systems, users typically have differing levels of access, depending on their security rating. Rather than assigning security levels to individual files, the system administrator can establish a policy that persons with particular security levels are to be denied access to documents on particular subjects. This simplifies the administration process, as the number of policies will typically be far less than the number of files on the network. This also allows for a document's content to change, thereby affecting the document's security rating, without the system administrator having to change the document's security level.

A person skilled in the art will recognize that content streams are not static. Content changes over time. For example, returning to the example of the newsgroup, threads die out as users stop posting new messages regarding the thread or moderators kill improper threads. New threads pop up as new subjects are proposed. People's viewpoints change as one argument or another sways them. As content changes, the need for policy enforcement can accordingly change. Since content streams are dynamic and change over time, it is expected that the distance between the content stream and the template will vary over time. Accordingly, impact summaries need to be updated to remain current. A person skilled in the art will recognize how FIG. 4 (and FIG. 5, below) can be modified to track changes in the content stream.

FIG. 5 shows a flowchart of a method according to an alternative embodiment of the invention to use a template to enforce network policy in the computer system of FIG. 1. The changes relative to the flowchart of the method in FIG. 4 lie in the middle of the method. Instead of monitoring all of the content stream and the entire network, at step 515 only a portion of the content stream is monitored to see how close it comes to the template. At step 520, a portion of the network is monitored to determine metadata about the content stream (e.g., the percentage of network traffic devoted to the content stream triggering the template). Finally, at step 527, data and metadata about the entire content stream is extrapolated from the sampled data. For example, if only 1/3 of the content stream is monitored, then the collected data is multiplied by a factor of 3 to characterize the entire content stream.

Having illustrated and described the principles of our invention in a preferred embodiment thereof, it should be readily apparent to those skilled in the art that the invention can be modified in arrangement and detail without departing from such principles. We claim all modifications coming within the spirit and scope of the accompanying claims.